Cancellation Analysis (April 2021)

Maxwell Knowles

4/20/2021

Seleced Reasons for Churn
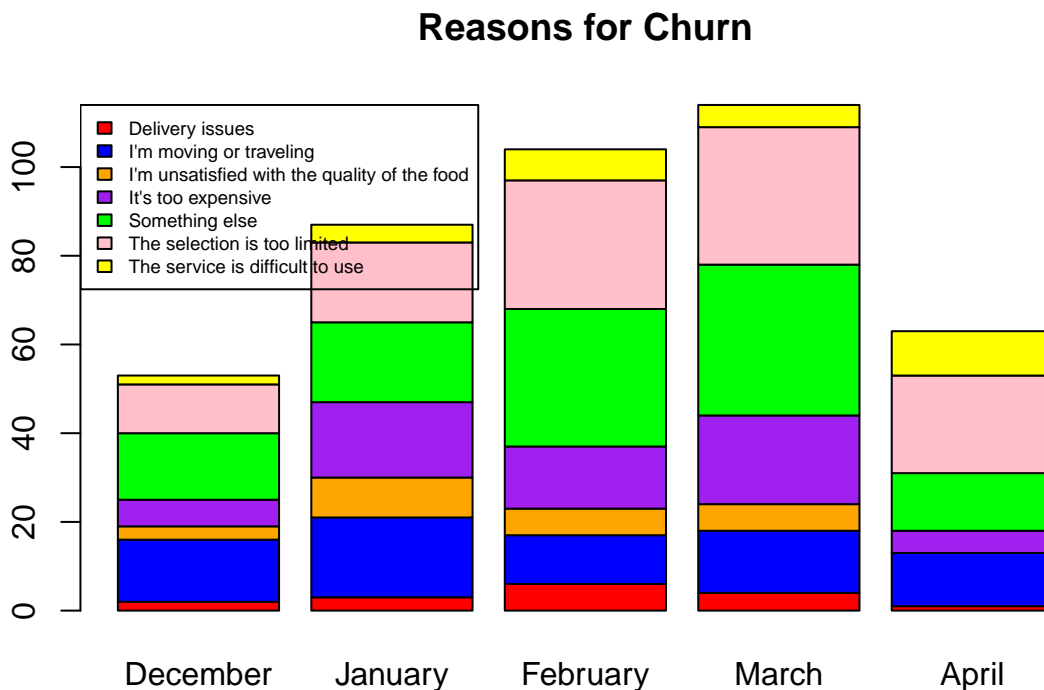
Reasons for Churn

The service being "difficult to use" has jumped as a share of selected reasons for churn in April, with "selection being too limited" growing as a selected reason over the last few months as well.

In aggregate, "The selection is too limited" and "Something else" stand as the top reasons for cancellation, followed by "I'm moving or traveling," "I'm unsatisfied with the quality of the food," "The service is difficult to use," "It's too expensive," and "Delivery issues." Churning over selection is a bit odd, since our selection is available for browsing and shopping before conversion. It's possible, however, that customers coming from other services just aren't thoroughly scanning the site and are ultimately disappointed by the narrower selection. Further, our search and category issues may be preventing customers from fully appreciating the SKUs we do in fact have (and continue to add).

When we regress reasons for churn against length of membership, we find a statistically significant relationship between "I'm unsatisfied with the quality of the food" and the duration of the (now canceled) membership (less surprising, there's also a statistically significant relationship between "moving or traveling" and membership duration). This means that when a customer's stated reason for canceling is food quality, the length of the membership was probably longer. This *may* imply that either there is a higher tolerance for lower food quality *or* that poor food quality is less common than the other issues.

Risk Range

Proportion Of Churn Within 3 Days of Monthly Payment
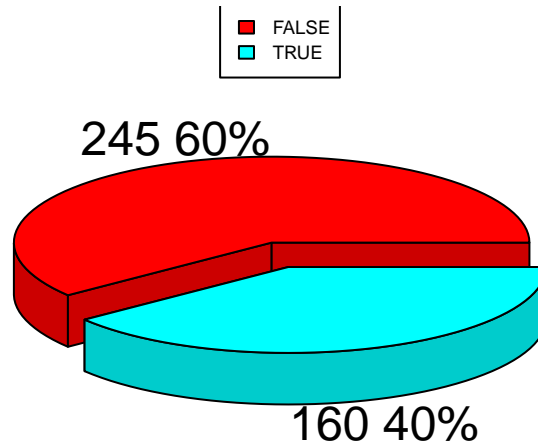
One interesting finding is that 40% of cancellations occur within a "risk range" of 3 days prior to 3 days post the ex-member's monthly subscription fee (we would expect **roughly** 20%). In this sample of customers, being in the risk range is **more common than any individual reason** selected from the exit survey.

Regressing reasons for churn against the binary risk range, we find the strongest statistically significant relationship to be between a customer in their risk range and "It's too expensive" as the selected reason for leaving.

Potential Next Steps

- Consider creating moments of delight leading up to a customer's monthly subscription payment

- Explore the relationship between ZenDesk tickets and stated reasons for churning to a) surface what our lead time might be from the point of infraction to the cancellation event, then b) devise manual, then automated ways of addressing these infractions

Regressions

```
##
## Call:
## glm(formula = Duration ~ Reason, family = "gaussian", data = df_gone)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -133.27   -61.24   -29.24    46.28   311.41
##
## Coefficients:
##                                                Estimate Std. Error t value
## (Intercept)                                      89.250     23.276   3.834
## ReasonI'm moving or traveling                    66.023     25.944   2.545
## ReasonI'm unsatisfied with the quality of the food  50.837   30.309   1.677
## ReasonIt's too expensive                         29.987     26.243   1.143
```

```
## ReasonSomething else                                            32.845     24.987    1.315
## ReasonThe selection is too limited                             -13.534     24.926   -0.543
## ReasonThe service is difficult to use                           -7.657     29.374   -0.261
##                                                       Pr(>|t|)
## (Intercept)                                           0.000146 ***
## ReasonI'm moving or traveling                         0.011311 *
## ReasonI'm unsatisfied with the quality of the food 0.094275 .
## ReasonIt's too expensive                              0.253859
## ReasonSomething else                                 0.189430
## ReasonThe selection is too limited                   0.587443
## ReasonThe service is difficult to use                0.794468
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 8668.363)
##
##     Null deviance: 3782225  on 404   degrees of freedom
## Residual deviance: 3450008  on 398   degrees of freedom
## AIC: 4830.6
##
## Number of Fisher Scoring iterations: 2


##
## Call:
## glm(formula = `Risk Range` ~ Reason, family = "binomial", data = df_gone)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q      Max
## -1.1918  -0.9794  -0.9107   1.3095   1.8297
##
## Coefficients:
##                                                    Estimate Std. Error z value
## (Intercept)                                         -1.4663     0.6405  -2.289
## ReasonI'm moving or traveling                        1.1610     0.6872   1.689
## ReasonI'm unsatisfied with the quality of the food   1.3793     0.7645   1.804
## ReasonIt's too expensive                             1.5002     0.6914   2.170
## ReasonSomething else                                 0.9808     0.6713   1.461
## ReasonThe selection is too limited                   0.8006     0.6717   1.192
## ReasonThe service is difficult to use                1.0916     0.7508   1.454
##                                                    Pr(>|z|)
## (Intercept)                                          0.0221 *
## ReasonI'm moving or traveling                        0.0912 .
## ReasonI'm unsatisfied with the quality of the food   0.0712 .
## ReasonIt's too expensive                             0.0300 *
## ReasonSomething else                                 0.1440
## ReasonThe selection is too limited                   0.2333
## ReasonThe service is difficult to use                0.1459
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 543.48  on 404   degrees of freedom
## Residual deviance: 534.75  on 398   degrees of freedom
```

```
## AIC: 548.75
##
## Number of Fisher Scoring iterations: 4
```